

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/103037/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Liu, Zihao, Liang, Gehao, Tan, Luyuan, Su, An, Jiang, Wen Guo ORCID: <https://orcid.org/0000-0002-3283-1111> and Gong, Chang 2017. High-efficient screening method for identification of key genes in breast cancer through microarray and bioinformatics. Anticancer Research 37 (8) , pp. 4329-4335. 10.21873/anticancer.11826 file

Publishers page: <http://dx.doi.org/10.21873/anticancer.11826>  
<<http://dx.doi.org/10.21873/anticancer.11826>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



- 12 Li QJ, Chau J, Ebert PJ, Sylvester G, Min H, Liu G, Braich R, Manoharan M, Soutschek J, Skare P, Klein LO, Davis MM and Chen CZ: miR-181a is an intrinsic modulator of T cell sensitivity and selection. *Cell* **129**: 147-161, 2007.
- 13 <https://www.mycancergenome.org/content/disease/lung-cancer/>.
- 14 Chen Y, Zhu X, Zhang X, Liu B and Huang L: Nanoparticles modified with tumor-targeting scFv deliver siRNA and miRNA for cancer therapy. *Molecular therapy: the journal of the American Society of Gene Therapy* **18**: 1650-1656, 2010.
- 15 Choi KY, Silvestre OF, Huang X, Hida N, Liu G, Ho DN, Lee S, Lee SW, Hong JI and Chen X: A nanoparticle formula for delivering siRNA or miRNAs to tumor cells in cell culture and *in vivo*. *Nature protocols* **9**: 1900-1915, 2014.
- 16 Hsu SH, Yu B, Wang X, Lu Y, Schmidt CR, Lee RJ, Lee LJ, Jacob ST and Ghoshal K: Cationic lipid nanoparticles for therapeutic delivery of siRNA and miRNA to murine liver tumor. *Nanomedicine: nanotechnology, biology, and medicine* **9**: 1169-1180, 2013.
- 17 Bethune G, Bethune D, Ridgway N and Xu Z: Epidermal growth factor receptor (EGFR) in lung cancer: an overview and update. *Journal of thoracic disease* **2**: 48-51, 2011.
- 18 Langenfeld E, Hong CC, Lanke G and Langenfeld J: Bone morphogenetic protein type I receptor antagonists decrease growth and induce cell death of lung cancer cell lines. *PLoS One* **8**: e61256, 2013.
- 19 Hao J, Lee R, Chang A, Fan J, Labib C, Parsa C, Orlando R, Andresen B and Huang Y: DMH1, a small molecule inhibitor of BMP type I receptors, suppresses growth and invasion of lung cancer. *PLoS One* **9**: e90748, 2014.

*Received May 2, 2017*

*Revised May 19, 2017*

*Accepted May 25, 2017*

# High-efficient Screening Method for Identification of Key Genes in Breast Cancer Through Microarray and Bioinformatics

ZIHAO LIU<sup>1\*</sup>, GEHAO LIANG<sup>1\*</sup>, LUYUAN TAN<sup>1</sup>, AN SU<sup>1</sup>, WENGUO JIANG<sup>2</sup> and CHANG GONG<sup>1,2</sup>

<sup>1</sup>Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetic and Gene Regulation, Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, P.R. China;

<sup>2</sup>Cardiff China Medical Research Collaborative, Cardiff University School of Medicine, Cardiff University, Cardiff, U.K.

**Abstract.** *Background/Aim:* The aim of the present study was to identify key pathways and genes in breast cancer and develop a new method for screening key genes with abnormal expression based on bioinformatics. *Materials and Methods:* Three microarray datasets GSE21422, GSE42568 and GSE45827 were downloaded from the Gene Expression Omnibus (GEO) database and differentially expressed genes (DEGs) were analyzed using GEO2R. The gene ontology (GO) and pathway enrichment analysis were established through DAVID database. The protein-protein interaction (PPI) network was performed through the Search Tool for the Retrieval of Interacting Genes (STRING) database and managed by Cytoscape. The overall survival (OS) analysis of the 4 genes including AURKA, CDH1, CDK1 and PPARG that had higher degrees in this network was uncovered Kaplan-Meier analysis. *Results:* A total of 811 DEGs were identified in breast cancer, which were enriched in biological processes, including cell cycle, mitosis, vessel development and lipid metabolic. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that the up-regulated DEGs were particularly involved in cell cycle, progesterone-mediated oocyte maturation and leukocyte transendothelial migration, while the down-regulated DEGs were mainly involved in regulation of lipolysis, fatty acid degradation and glycerolipid metabolism. Through PPI network analysis, 14

hub genes were identified. Among them, the high expression of AURKA, CDH1 and CDK1 were associated with worse OS of breast cancer patients; while the high expression of PPARG was linked with better OS. *Conclusion:* The present study identified key pathways and genes involved in breast cancer which are potential molecular targets for breast cancer treatment and diagnosis.

Breast cancer, originating from the epithelium of the mammary gland, has undergone an increase in incidence in China in the last few decades and remains the most frequently diagnosed cancer among women, with an estimated 272,400 new cases and 70,700 related deaths in 2015 (1, 2). Despite advances in endocrine and surgical therapy as well as chemotherapy, breast cancer-associated mortality in China still increases (2). This is attributed, at least partly, to the difficulties in breast cancer diagnosis and lack of treatments for advanced breast cancer patients.

Although immunohistochemistry analysis allows distribution of breast cancer tumors into three major subgroups: hormone receptor (HR)-positive, human epidermal receptor 2 (HER2)-positive and triple-negative breast cancer (3). Each subgroup shows specific prognosis and different responses to anticancer therapy, there is, however, accumulating evidence supporting the hypothesis that they share similar activated genes as well as signaling pathways (4, 5). These signaling pathways and multiple genes participate in the progression of breast cancer. Therefore, there is an urgent need to identify the key genes involved in breast cancer and explore highly efficient key-gene screening methods.

In addition, breast cancer, which is characterized by cumulative epigenetic and genetic aberrations as well as cancer cell heterogeneity, is certainly a complex disease. Understanding the pathological molecular alternations governing breast cancer initiation and progression is greatly important for prognosis and cancer prevention.

\*These Authors contributed equally to this work.

**Correspondence to:** Chang Gong, Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, 107 Yanjiang West Road, Guangzhou, 510120, P.R. China. Tel: +86 208133783, Fax: +86 2081332853, e-mail: changgong282@163.com

**Key Words:** Breast cancer, bioinformatics analysis, differentially expressed genes.

Since the introduction of bioinformatics, such as gene microarray technology and high-throughput sequence, which detect genetic alternations, we are able to analyze gene expression profiles during carcinogenesis and cancer progression. It can also uncover thousands of differentially expressed genes (DEGs) associated with different biological processes, different pathways, or molecular functions of cancer. Moreover, the bioinformatics methods make it possible to comprehensively analyze large amounts of data from the microarray output. Given the false-positives and heterogeneity of different microarray results, we processed three microarray datasets to obtain DEGs between normal and breast cancer tissues. Combined with bioinformatics, we identified key pathways and genes in breast cancer. Through survival analysis we found that *PPARG*, *AURKA*, *CDH1* and *CDK1* were critical DEGs in breast cancer.

## Materials and Methods

**Microarray data and data processing.** Three gene expression profiles (GSE21422, GSE42568 and GSE45827) were obtained from public microarray data storage platforms (the Gene Expression Omnibus, GEO, <http://www.ncbi.nlm.nih.gov/geo/>). GSE21422 included 5 normal breast tissue samples and 14 breast cancer samples (6). GSE42568 contained 17 normal breast tissue samples and 104 breast cancer samples (7). GSE45827 consisted of 11 normal breast tissue samples and 130 breast cancer samples (8). GEO2R (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>) was applied to screen DEGs between normal breast tissue and breast cancer in the three datasets. The *p*-value was adjusted for the correction of false positive outputs when using the Benjamini and Hochberg (False discovery rate) method and logFC represents the fold changes of down- or up-regulated genes of breast cancer against normal breast tissues. According to other studies, we set the adjusted *p*-value<0.01 and |logFC|>1 as significant DEGs (9, 10).

**Functional and pathway enrichment analysis of DEGs.** By GO enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis method, we mapped the potential relevant biological function annotation of the significant DEGs comprehensively. GO and KEGG was applied through DAVID database (<https://david.ncifcrf.gov/>) which is an online tool for gene annotation, function visualization and large volume data integration (11). A *p*-value <0.05 was considered statistically significant.

**The protein-protein interaction (PPI) network and module selection.** To evaluate the PPI network information of the significant DEGs, we used an online database, Search Tool for the Retrieval of Interacting Genes (STRING, <http://www.string-db.org/>). Since we analyze more than 10,000 DEGs, we set confident interaction score  $\geq 0.7$  (high confident) as significant interaction in order to avoid uncertain/unidentified PPI. The PPI networks were visualized *via* Cytoscape software and modules of PPI network were screened using Molecular Complex Detection (MCODE) with the included criteria as follow: degree cutoff=2, node score cutoff=0.2, k-core=2, max. depth=100 (9, 10). The average degrees of MCODE score and nodes in modules were chosen as threshold, thus we set MCODE

scores  $\geq 9$  and hub nodes  $\geq 9$  as criteria. The functional enrichment analysis was performed through DAVID in the modules.

**Survival analysis of hub genes in breast cancer.** According to the expression of hub genes, breast cancer patients were divided into two groups (high and low). The overall survival of the two groups was assessed by Kaplan-Meier plotter ([www.kmplot.com](http://www.kmplot.com)) while the log rank *p*-value and the hazard ratio (HR) (95% confidence intervals) were calculated.

## Results

**Identification of DEGs.** A total of 875, 1,324 and 4,144 DEGs were up-regulated in GSE21422, GSE42568 and GSE45827 datasets, respectively (Figure 1A); while 1,145, 1,551 and 1,726 DEGs were down-regulated in GSE21422, GSE42568 and GSE45827 datasets (Figure 1B). In total, 811 DEGs presented the same expression tendency in all three datasets, including 344 up-regulated and 467 down-regulated in breast cancer tissues compared to normal breast tissues.

**Gene ontology (GO) and pathway enrichment analysis.** Through DAVID tool, we analyzed the GO and KEGG pathway enrichment of the identified DEGs (Table I). GO biological process (BP) analysis indicated the up-regulated DEGs were mainly involved in those regulating mitotic cell cycle and cell cycle; the down-regulated DEGs were mainly enriched in those relating to vessel development and lipid metabolic process. For GO cell component (CC), the up-regulated DEGs were mainly enriched in genes in microtubule cytoskeleton, chromosome, the centromeric region and spindle; the down-regulated DEGs were mainly enriched in those related to lipid particle, membrane raft and membrane microdomain. Concerning molecular function, the up-regulated DEGs were mainly enriched in microtubule binding, cytoskeletal protein binding and enzyme binding; the down-regulated DEGs were mainly enriched in glycosaminoglycan binding, hormone binding and RNA polymerase II core promoter proximal region sequence-specific binding. In addition, KEGG analysis showed that the up-regulated DEGs were mainly enriched in cell cycle, progesterone-mediated oocyte maturation and leukocyte transendothelial migration signaling pathways and the down-regulated DEGs were mainly enriched in regulation of lipolysis in adipocytes, fatty acid degradation and glycerolipid metabolism (Table I).

**PPI network and module selection.** The PPI network consisted of 443 nodes and 1181 edges (Figure 2A). Based on the STRING output, there are 57 genes with their degree  $\geq 10$  and 192 genes with their module score  $\geq 1$ , respectively. In order to select the most significant genes, we set criteria as module score >1 and degree >20. Twelve hub genes including: *IGF1*, *LEP*, *KIF11*, *PTEN*, *FOXO1*, *FGF2*,



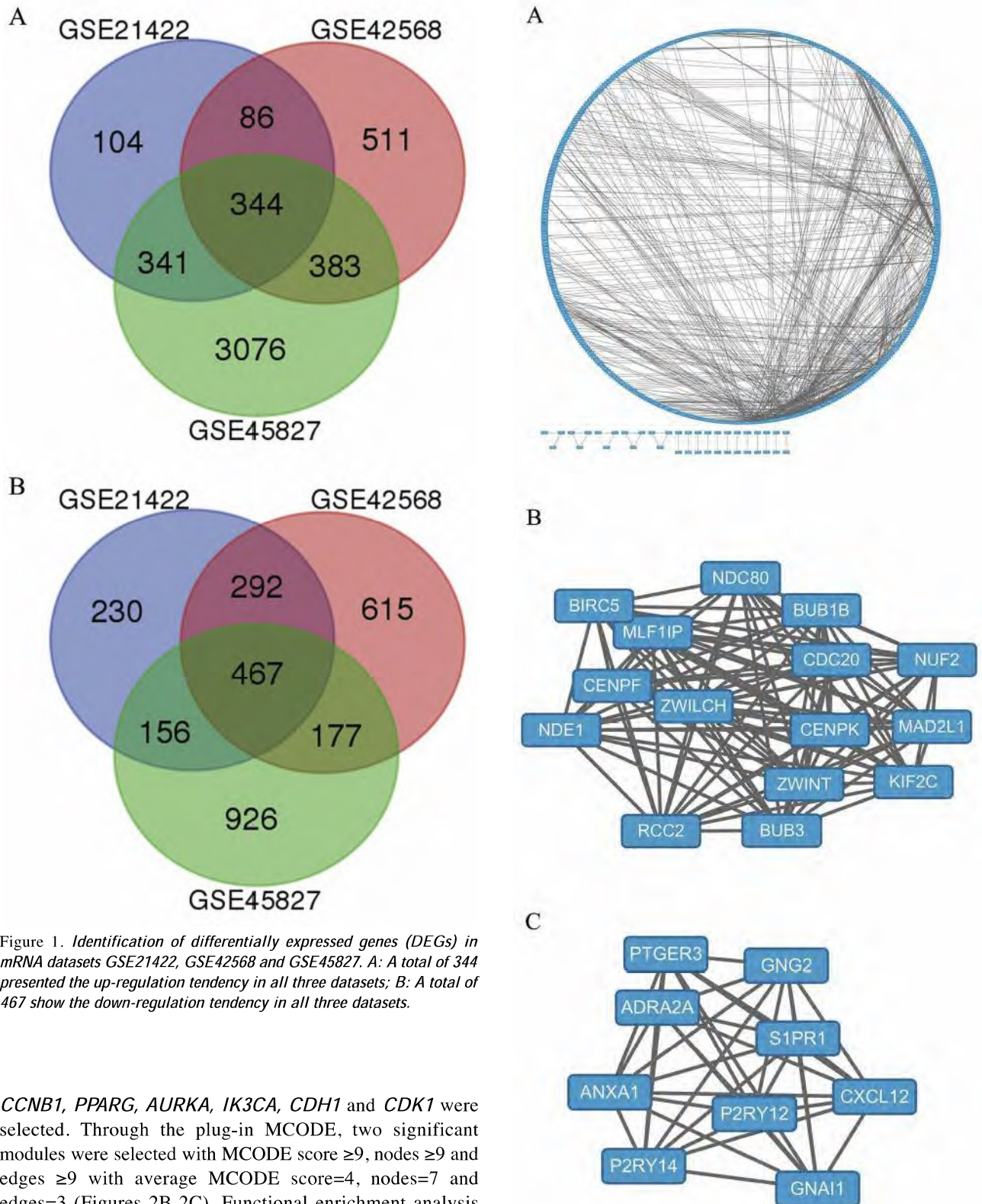


Figure 1. Identification of differentially expressed genes (DEGs) in mRNA datasets GSE21422, GSE42568 and GSE45827. A: A total of 344 presented the up-regulation tendency in all three datasets; B: A total of 467 show the down-regulation tendency in all three datasets.

*CCNB1*, *PPARG*, *AURKA*, *IK3CA*, *CDH1* and *CDK1* were selected. Through the plug-in MCODE, two significant modules were selected with MCODE score  $\geq 9$ , nodes  $\geq 9$  and edges  $\geq 9$  with average MCODE score=4, nodes=7 and edges=3 (Figures 2B-2C). Functional enrichment analysis indicated that up-regulated genes in module one were significantly enriched in cell division and mitotic nuclear division, while the down-regulated genes in module two

Figure 2. Protein-protein interaction (PPI) network and selected modules with a higher degree. A: The constructed PPI network of differentially expressed genes; B: Module 1; C: Module 2.

Table I. Gene ontology and pathway enrichment analysis of differentially expressed genes in breast cancer.

Category	Term	Gene function	Gene count	p-Value
Up-regulated				
GOTERM_BP_FAT	GO:0000278	Mitotic cell cycle	79	1.20E-28
GOTERM_BP_FAT	GO:1903047	Mitotic cell cycle process	75	5.60E-28
GOTERM_BP_FAT	GO:0022402	Cell cycle process	92	1.10E-27
GOTERM_CC_FAT	GO:0015630	Microtubule cytoskeleton	65	4.20E-14
GOTERM_CC_FAT	GO:0000775	Chromosome, centromeric region	25	1.50E-12
GOTERM_CC_FAT	GO:0005819	Spindle	31	1.60E-12
GOTERM_MF_FAT	GO:0008017	Microtubule binding	16	1.50E-05
GOTERM_MF_FAT	GO:0008092	Cytoskeletal protein binding	34	4.70E-05
GOTERM_MF_FAT	GO:0019899	Enzyme binding	57	6.00E-05
KEGG_PATHWAY	hsa04110	Cell cycle	15	8.70E-08
KEGG_PATHWAY	hsa04914	Progesterone-mediated oocyte maturation	7	6.40E-03
KEGG_PATHWAY	hsa04670	Leukocyte transendothelial migration	8	7.50E-03
Down-regulated				
GOTERM_BP_FAT	GO:0001568	Lipid transport	52	1.60E-15
GOTERM_BP_FAT	GO:0001944	Vasculature development	53	4.20E-15
GOTERM_BP_FAT	GO:0006629	Lipid metabolic process	82	1.40E-14
GOTERM_CC_FAT	GO:0005811	Lipid particle	13	1.50E-07
GOTERM_CC_FAT	GO:0045121	Membrane raft	25	4.20E-07
GOTERM_CC_FAT	GO:0098857	Membrane microdomain	25	4.50E-07
GOTERM_MF_FAT	GO:0005539	Glycosaminoglycan binding	20	1.10E-06
GOTERM_MF_FAT	GO:0042562	Hormone binding	10	2.70E-05
GOTERM_MF_FAT	GO:0000982	Transcription factor activity, RNA polymerase II core Promoter proximal region sequence-specific binding	23	4.90E-05
KEGG_PATHWAY	hsa04923	Regulation of lipolysis in adipocytes	14	6.20E-09
KEGG_PATHWAY	hsa00071	Fatty acid degradation	9	3.80E-05
KEGG_PATHWAY	hsa00561	Glycerolipid metabolism	10	4.70E-05

If there were more than four terms enriched in this category, top four terms were selected according to p-value. Count: the number of enriched genes in each term.

were mainly enriched in positive regulation of wound healing and inflammatory response.

**Hub genes expression analysis and Kaplan-Meier plotter.** We picked out 3 significant hub genes from the 12 hub genes with their expression more than two-fold higher than normal tissue and 1 hub gene with its expression more than two-fold lower than normal tissue in all three datasets (GSE21422, GSE42568 and GSE45827). Through Kaplan-Meier plotter, we found high expression of *PPARG* (HR 0.66 [0.53-0.82],  $p < 0.01$ ) which was associated with better overall survival for breast cancer patients. On the other side, high expression of *AURKA* (HR 1.83 [1.47-2.28],  $p < 0.01$ ), *CDH1* (HR 1.39 [1.12-1.72],  $p < 0.01$ ) and *CDK1* (HR 1.55 [1.39-1.73],  $p < 0.01$ ) were associated with worse overall survival for breast cancer patients (Figure 3A-3D).

## Discussion

In the present study, a total of 33 normal breast tissue samples and 248 breast cancer tissues were retrieved from the GEO

database. A total of 811 DEGs including 344 up-regulated and 467 down-regulated genes were screened. In order to demonstrate the molecular function and the protein-protein interactions of these abnormal DEGs, we performed GO and KEGG pathway analysis. GO terms and KEGG analysis indicated that the up-regulated DEGs were mainly enriched in cell cycle, microtubule cytoskeleton and mitosis. In line with this, the abnormal function of cell cycle and mitosis would appear to be a major cause of cancer (12, 13). The down-regulated DEGs were mainly enriched in hormone binding, lipid metabolism, and glycosaminoglycan binding. In our study we found some lipid metabolism genes that were down-regulated. A recent work has also reported that some genes involved in lipid metabolism are down-regulated in cancer and potentially function as tumor suppressors (14). In addition, the down-regulated DEGs were mainly enriched in hormone binding. Strikingly, previous research has revealed that breast cancer cells can transform from HER2-positive to HER2-negative and become chemotherapy resistance (15). Thus, breast cancer potentially down-regulates its hormone-binding ability to become endocrine therapy resistant.

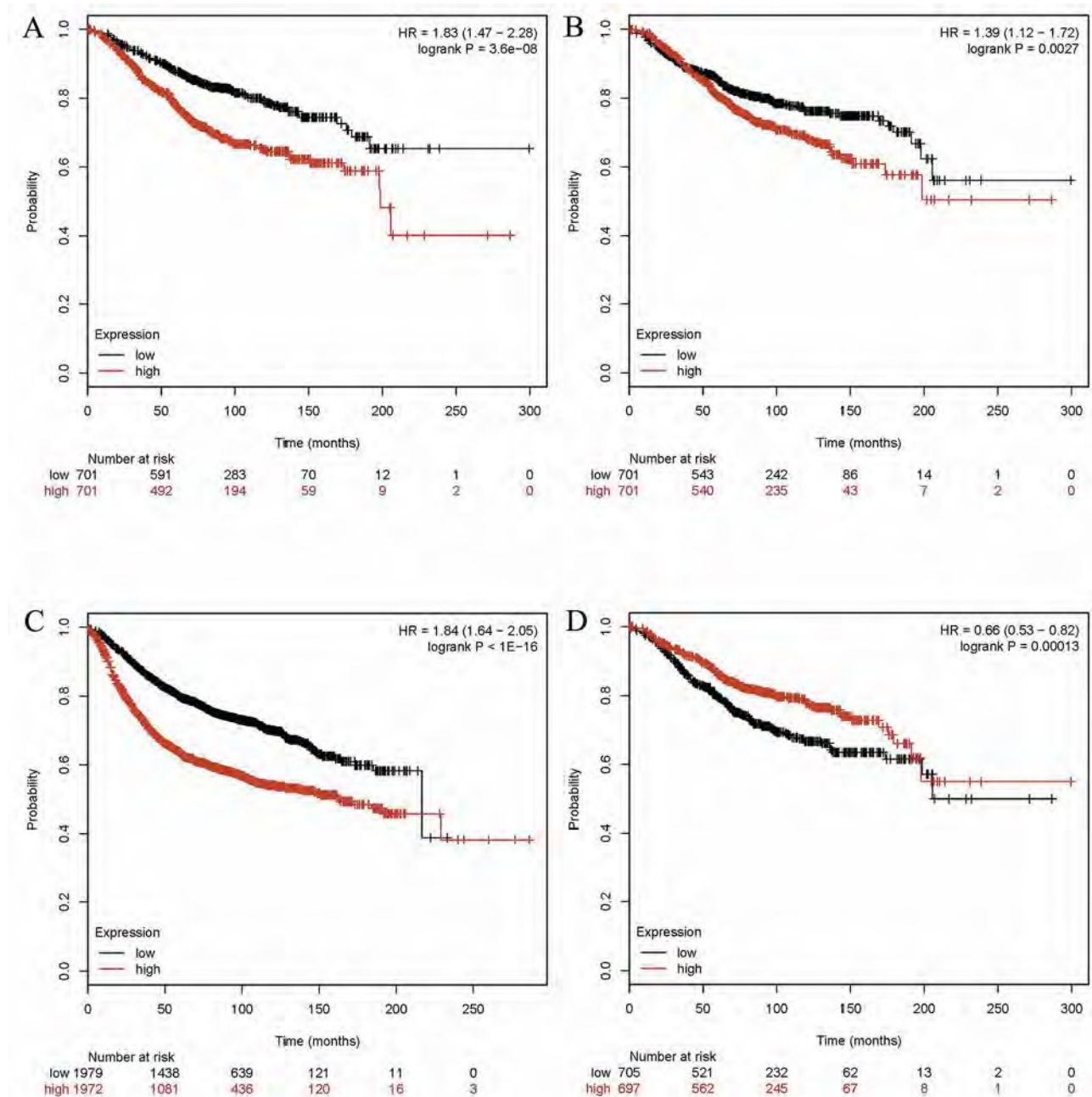


Figure 3. Kaplan-Meier plotter of four genes in breast cancer. A: AURKA (Affymetrix ID: 204092\_s\_at); B: CDH1 (Affymetrix ID: 201131\_s\_at); C: CDK1 (Affymetrix ID: 210559\_s\_at); D: PPARG (Affymetrix ID: 208510\_s\_at). HR: Hazard ratio; CI: confidence interval.

The combination of bioinformatics and microarray has enabled us to study a vast number of genes with abnormal expression and has improved the efficiency of analyzing complex genetic diseases such as cancer by more than five orders of magnitude since 2005. However, the occurrence of false-positive and false-negative results still exists (16). In order to avoid false-positive/negative and the

heterogeneity between different sequencing platforms, we combined three microarray datasets together. We finally identified 4 hub genes which were found to be expressed more than two-fold higher/ lower in breast cancer tissues in all 3 GEO datasets. Survival analysis indicated high expression of *PPARG* was associated with better overall survival of breast cancer patients and the upregulation of

*AURKA*, *CDH1* and *CDK1* were significantly associated with worse overall survival of patients. Thus, molecular drugs targeting these 4 key genes might potentially be applicable in all breast cancer patients.

*AURKA* is a key regulator of mitosis, especially important in the passage from G<sub>2</sub> to M (17). The forced expression of *AURKA* will result in gene copy number alternations and genetic mutations which are associated with the emergence of breast carcinoma (18). Further, *AURKA* overexpression correlates with high risk of breast cancer (19). Consistent with this, we found *AURKA* highly expressed in breast cancer patients and a predictor of poor prognosis. This indicates *AURKA* may function in breast tumor initiation and progression. The second hub gene, *CDH1*, encodes the E-cadherin protein. Concrete evidence, strikingly, supports that E-cadherin functions as a tumor promoter. Clinical specimens have shown high expression of E-cadherin in 12% of lobular carcinomas and 70% of invasive ductal carcinomas (20, 21). Our study also found that high levels of *CDH1* correlate with worse prognosis. Interestingly, the soluble form of E-cadherin contributes to tumor progression and metastasis *via* undermining cell junctions and activates onco-pathways (22). *CDK1*, belonging to the serine/threonine kinase family, is directly involved in regulation of cell cycle. It is essential to the initiation of mitosis and transition process of the cell cycle; the loss of function of *CDK1* will lead to G<sub>2</sub> phase cycle arrest (23, 24). *PPARG* is a potential tumor suppressor and is down-regulated in breast cancer tissue. Overexpression of *PPARG* associates with better prognosis for patients. *PPARG* is a member of the nuclear steroid receptor superfamily. *In vitro*, *PPARG* activation will inhibit breast cancer cell proliferation, migration and invasion (25). Gene meta-analysis and clinical data suggest that a high level of *PPARG* is a protective factor for breast cancer and colorectal cancer, which is consistent with our findings (26, 27).

In summary, the present study provided a comprehensive analysis of DEGs through bioinformatics to find potential genes that are involved in breast cancer progression and suggest useful targets for investigation or new biomarkers for cancer diagnosis.

## Acknowledgements

This work was supported by grants from the National Key Research and Development Program of China (2017YFC1309100); the Natural Science Foundation of China (81472466 and 81672594); National Science Foundation of Guangdong Province (2014A03036003, 2014A030310378, 2014A020212059, 2015A030313172, 2015B050501004, 2016A050502018 and 2016A030313237); China Scholarship Council (No. 201606385034); Cultivation for Major Projects and Emerging Interdisciplinary Funding Project of Sun Yat-sen University (17ykjc13) and Elite Young Scholars Program of Sun Yat-sen Memorial Hospital (Y201401). The Authors would also like to thank Cancer Research Wales and Wales National Life Science Research Network (NRN), Cardiff, Wales for their support.

## References

- Videira M, Reis RL and Brito MA: Deconstructing breast cancer cell biology and the mechanisms of multidrug resistance. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **1846**: 312-325, 2014.
- Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J: Cancer statistics in China, 2015. *CA: a cancer j clin* **66**: 115-132, 2016.
- Prat A and Perou CM: Deconstructing the molecular portraits of breast cancer. *Mol Oncol* **5**: 5-23, 2011.
- Adamo B, Deal AM, Burrows E, Geradts J, Hamilton E, Blackwell KL, Livasy C, Fritchie K, Prat A and Harrell JC: Phosphatidylinositol 3-kinase pathway activation in breast cancer brain metastases. *Breast Cancer Res* **13**: R125, 2011.
- Nandini D, Pradip and Brian LJ: PI3K-AKT-mTOR inhibitors in breast cancers: From tumor cell signaling to clinical trials. *Pharmacol Ther*, 2017. doi: 10.1016/j.pharmthera.2017.02.037. [Epub ahead of print]
- Kretschmer C, Sterner-Kock A, Siedentopf F, Schoenegg W, Schlag PM and Kemmner W: Identification of early molecular markers for breast cancer. *Mol Cancer* **10**: 15, 2011.
- Clarke C, Madden SF, Doolan P, Aherne ST, Joyce H, O'Driscoll L, Gallagher WM, Hennessy BT, Moriarty M and Crown J: Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis* **34**: 2300-2308, 2013.
- Gruosso T, Mieulet V, Cardon M, Bourachot B, Kieffer Y, Devun F, Dubois T, Dutreix M, Vincent-Salomon A and Miller KM: Chronic oxidative stress promotes H2AX protein degradation and enhances chemosensitivity in breast cancer patients. *Embo Mol Med* **8**: 527-549, 2016.
- Xu Z, Zhou Y, Cao Y, Dinh TLA, Wan J and Zhao M: Identification of candidate biomarkers and analysis of prognostic values in ovarian cancer by integrated bioinformatics analysis. *Med Oncol* **33**: 130, 2016.
- Liang B, Li C and Zhao J: Identification of key pathways and genes in colorectal cancer using bioinformatics analysis. *Med Oncol* **33**: 111, 2016.
- Huang DW, Sherman BT and Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocol* **4**: 44-57, 2009.
- Kastan MB and Bartek J: Cell-cycle checkpoints and cancer. *Nature* **432**: 316, 2004.
- Malumbres M and Barbacid M: Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* **9**: 153-166, 2009.
- Keckesova Z, Donaher JL, Cock JD, Freinkman E, Lingrell S, Bachovchin DA, Bieri B, Tischler V, Noske A and Okondo MC: LACTB is a tumour suppressor that modulates lipid metabolism and cell state. *Nature* **543**: 681-686, 2017.
- Jordan NV, Bardia A, Wittner BS, Benes C, Ligorio M, Zheng Y, Yu M, Sundaresan TK, Licausi JA and Desai R: HER2 expression identifies dynamic functional states within circulating breast cancer cells. *Nature* **537**: 102-106, 2016.
- Pounds S and Morris SW: Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values. *Bioinformatics* **19**: 1236-1242, 2003.
- Nigg EA: Centrosome aberrations: cause or consequence of cancer progression? *Nat Rev Cancer* **2**: 815, 2002.



- 18 Treekitkarnmongkol W, Katayama H, Kai K, Sasai K, Jones JC, Wang J, Shen L, Sahin AA, Gagea M and Ueno NT: Aurora kinase-A overexpression in mouse mammary epithelium induces mammary adenocarcinomas harboring genetic alterations shared with human breast cancer. *Carcinogenesis* **37**: 1180-1189, 2016.
- 19 Cox DG, Hankinson SE and Hunter DJ: Polymorphisms of the AURKA (STK15/Aurora Kinase) Gene and Breast Cancer Risk (United States). *Cancer Cause Control* **17**: 81-83, 2006.
- 20 Christiansen JJ and Rajasekaran AK: Reassessing Epithelial to Mesenchymal Transition as a Prerequisite for Carcinoma Invasion and Metastasis. *Cancer Res* **66**: 8319-8326, 2006.
- 21 Kowalski PJ, Rubin MA and Kleer CG: E-cadherin expression in primary carcinomas of the breast and its distant metastases. *Breast Cancer Res* **5**: 217-222, 2003.
- 22 Hu QP, Kuang JY, Yang QK, Bian XW and Yu SC: Beyond a tumor suppressor: Soluble E-cadherin promotes the progression of cancer. *Int J Cancer* **138**: 2804, 2015.
- 23 Shapiro GI: Cyclin-Dependent Kinase Pathways As Targets for Cancer Treatment. *J Clin Oncol* **24**: 1770-1783, 2006.
- 24 Malumbres M: Cyclin-dependent kinases. *Genome Biol* **101**: 2511-2526, 2014.
- 25 Zaytseva YY, Wallis NK, Southard RC and Kilgore MW: The PPAR $\gamma$  antagonist T0070907 suppresses breast cancer cell proliferation and motility *via* both PPAR $\gamma$ -dependent and -independent mechanisms. *Anticancer Res* **31**: 813-823, 2011.
- 26 Tang W, Chen Y, Wang Y, Gu H, Chen S and Kang M: Peroxisome proliferator-activated receptor gamma (PPARG) polymorphisms and breast cancer susceptibility: a meta-analysis. *Int J Clin Exp Med* **8**: 12226-12238, 2015.
- 27 Ogino S, Shima K, Baba Y, Nosho K, Irahara N, Kure S, Chen L, Toyoda S, Kirkner GJ and Wang YL: Colorectal cancer expression of peroxisome proliferator-activated receptor gamma (PPARG, PPARGgamma) is associated with good prognosis. *Gastroenterology* **136**: 1242-1250, 2009.

*Received May 2, 2017*

*Revised May 21, 2017*

*Accepted May 22, 2017*